AFRL-IF-RS-TR-2003-231
Final Technical Report
October 2003

THEORETIC BASED PERFORMANCE ANALYSIS OF DISTRIBUTED SENSOR NETWORK

HRL Laboratories, LLC

Sponsored by
Defense Advanced Research Projects Agency
DARPA Order No. M271

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

AIR FORCE RESEARCH LABORATORY
INFORMATION DIRECTORATE
ROME RESEARCH SITE
ROME, NEW YORK

This report has been reviewed by the Air Force Research Laboratory, Information Directorate, Public Affairs Office (IFOIPA) and is releasable to the National Technical Information Service (NTIS). At NTIS it will be releasable to the general public, including foreign nations.

AFRL-IF-RS-TR-2003-231 has been reviewed and is approved for publication.

APPROVED: /s/
RICHARD C. BUTLER II
Project Engineer

FOR THE DIRECTOR: /s/
WARREN H. DEBANY, JR., Technical Advisor
Information Grid Division
Information Directorate

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	OCTOBER 2003	Final Sep 01 – May 03

4. TITLE AND SUBTITLE
THEORETIC BASED PERFORMANCE ANALYSIS OF DISTRIBUTED SENSOR NETWORK

5. FUNDING NUMBERS
C - F30602-01-C-0192
PE - 62301E
PR - M271
TA - SE
WU - NS

6. AUTHOR(S)
Shubha Kadambe and Cindy Daniell

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
HRL Laboratories, LLC
3011 Malibu Canyon Road
Malibu California 90265

8. PERFORMING ORGANIZATION REPORT NUMBER
N/A

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Defense Advanced Research Projects Agency
3701 North Fairfax Drive
Arlington Virginia 22203-1714

AFRL/IFGA
525 Brooks Road
Rome New York 13441-4505

10. SPONSORING / MONITORING AGENCY REPORT NUMBER
AFRL-IF-RS-TR-2003-231

11. SUPPLEMENTARY NOTES

AFRL Project Engineer: Richard C. Butler II/IFGA/(315) 330-1888/ Richard.Butler@rl.af.mil

12a. DISTRIBUTION / AVAILABILITY STATEMENT	12b. DISTRIBUTION CODE
APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.	

13. ABSTRACT ***(Maximum 200 Words)***
Spatially distributed networks of inexpensive small and smart nodes, with multiple onboard sensors is an important class of emerging networked systems for various applications. Since this network of sensors has to operate efficiently in adverse environments, it is important that these sensors process information efficiently and share information such that the decision accuracy is improved. One way to address this problem is to measure the value of information obtained from multiple sensors on the same sensor node as well as from the neighboring nodes and fuse that information if value is added in terms of improvement in decision accuracy. In this study, information theoretic based and other measures are developed for assessing the value of information. These measures are then used to evaluate whether the value is added or not. The value is measured in terms of improvement in decision accuracy. The decisions that are considered in this study are target detection and classification. The value of information measures are then integrated with these two algorithms. These measures and value added are verified by using real data - SITEX02 and in the context of target detection and classification. From the results of improvement in classification accuracy and probability of detection, it can be seen that the utilization of measure of value of information while fusing helps in improving the decision accuracy significantly. This is an important finding in that the data from different sensors and/or from different nodes in a network of sensors can be efficiently made use of by using the measures developed here. This helps in efficient sensor management/selection. This has a significant impact in defense applicants such as situation awareness while generating a single integrated picture.

14. SUBJECT TERMS	15. NUMBER OF PAGES
Distributed Sensor Networks, Collaborative Processing, Data Fusion, Information Theoretic Based Metric	40
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

Table of Contents

List of Figures and Tables

1. Chapter: Introduction

This document constitutes the final report of contract F30602-01-C-0192, Information Theoretic Based Performance Analysis of Distributed Sensor Network. All work performed by HRL Laboratories, LLC under this contract since the beginning of the contract (18 September 2001) is recounted herein.

1.1. Contract Objectives:

As part of the SensIT August 02 experiment at Twentynine Palms, CA, a time series database was created. This database consists of time-series data of seismic, acoustic and PIR sensors sensing various activities. Under the "Collaborative signal processing (CSIP)" part of the SensIT program, several detection, tracking and classification algorithms have been developed and have been tested using this data. However, there is a need for theoretical analysis of performance of these algorithms, to determine the lower and upper bound of accuracy of detection, tracking and classification. This will illustrate when to use which algorithms. Therefore, in this study, we proposed to conduct such an analysis using information theoretic based or other value of information based metrics. For this analysis, we considered the SITEX 02 data. In particular, the main objectives of this study are as follows:

1 Develop information theoretic based and other metrics, to assess the value of information obtained from different sensors on the same sensor node and from the neighboring nodes.

2 Apply these metrics to assess the information obtained from the other sensors on the same node and/or from the neighboring nodes, and make a decision of when to or when not to fuse information.

3 Analyze the detection and classification algorithms developed as part of the CSIP of the SensIT program, using SITEX 02 data and these metrics.

These objectives were successfully met at the completion of this study and specific achievements of this study are listed in the next section.

1.2. Summary of achievements of this contract:

1. Measures based on Euclidean distance, Correlation, Mutual information and relative entropy (Kullback-Liebler) were developed.

2. An energy based detector and a maximum likelihood based classifier were considered.

3. The above mentioned measures were used in assessing the value of information obtained from multiple sensors on the same node and from the neighboring nodes. The value of information was in terms of improvement in decision accuracy, which corresponds to detection or classification accuracy in the context of a detector or a classifier. If the information obtained from the other sensors of the same node or from the neighboring nodes did not improve the decision accuracy, then it was discounted. Otherwise, it was fused with the existing information.

4. The measures and the performance of the energy based detector and the maximum likelihood based classifier were evaluated using the SITEX02 data.

5. The performance analysis included using (a) only one sensor on each node, with and without fusing information from the neighboring nodes, based on whether value was added or not, and with and without rejection capability; the performance without fusion corresponds to lower bound, (b) multiple sensors on each node, with and without fusing information from the other sensors, based on whether value was added or not and with and without the capability to reject the bad data, and (c) multiple sensors on each node, and the neighboring nodes with and without fusion and rejection capability; the performance with fusion of information from multiple sensors and the neighboring nodes when value is added and with the capability to reject the bad data corresponds to the upper bound.

6. The experimental results indicated that by fusing information from multiple sensors on the same node and from the neighboring nodes, when value was added, and by having the capability to reject bad information, the average performance of both a detector and a classifier improved significantly, and the upper bound of close to 100 % decision accuracy can be obtained.

7. An invention disclosure based on this study has been submitted. Also a technical paper was published in the proceedings of the Fusion'03 conference, held in Cairns, Australia from July 8-11, 2003

.

1.3. Outline of this report:

This final report is organized as follows:

- In chapter 2, the details of measures of value of information that we have developed under this study, are provided.
- In chapter 3, an energy based detector and a maximum likelihood based classifier that were considered in this study are briefly reviewed.
- In chapter 4, a brief description of the data, the experimental details, and the results are provided.
- In Chapter 5, we conclude and indicate the future directions.
- In Appendix A, a copy of the technical paper that is being published based on this study is included.

2. Chapter: Measures of value of information

A spatially distributed network of inexpensive, small and smart nodes with multiple onboard sensors is an important class of emerging networked systems for various defense and commercial applications. Since this network of sensors has to operate efficiently in adverse environments, using limited battery power and resources, it is important that these sensors process information efficiently and share information such that the decision accuracy is improved. In this paper, this is addressed by developing measures that assess the value of information obtained from multiple sensors on board a node and from the neighboring nodes, by conditioning it on improvement in the decision accuracy. If the information obtained from other sensor types on a node and/or from the neighboring nodes do improve the decision accuracy, then the information is fused. In our study, information is obtained in the form of features (for classification) or data (for detection). In [1-2], we developed a general information theoretic based metric that can be used in any kind of sensor selection and data fusion. However, while analyzing the real data with respect to a classifier and a detector, we observed that the correlation between the metric of value of information and the decision accuracy depends on the type of a classifier or a detector. Hence, we think that a mutual information metric may not always work. Therefore, we have developed several measures and studied them systematically with respect to one type of classifier and a detector. The mathematical details of these measures are provided in the following sections.

2.1. Mutual Information

Entropy is a measure of uncertainty. Let $H(x)$ be the entropy of previously observed x events. Let y be a new event. We can measure the uncertainty of x after including y by using the conditional entropy which is defined as:

$$H(x|y) = H(x,y) - H(y) \tag{2.1}$$

with the property $0 \le H(x|y) \le H(x)$. The conditional entropy $H(x|y)$ represents the amount of uncertainty remaining about x after y has been observed. If the uncertainty is reduced, then there is information gained by observing y. Therefore, we can measure the value of y by using conditional entropy. Another measure that is related to conditional entropy that one can use is the mutual information $I(x,y)$ which is a measure of uncertainty that is resolved by observing y and is defined as:

$$I(x,y) = H(x) - H(x|y). \tag{2.2}$$

To explain how this measure can be used to determine the value of information obtained from another sensor type, an example is provided below.

2.1.1. An example of value of information using mutual information as a metric

Let $A = \{a_k\}$ $k = 1, 2,\ldots$ be the set of features from sensor 1 and let $B = \{b_l\}$ $l = 1, 2,\ldots$ be the set of features from sensor 2 on the same node. Let $p(a_i)$ be the probability of feature a_i. Let $H(A)$, $H(B)$ and $H(A|B)$ be the entropy corresponding to sensor 1, sensor 2 and sensor 1 given sensor 2, respectively.They are defined as [3]:

$$\begin{aligned} H(A) &= \sum_k p(a_k)\log\left(\frac{1}{p(a_k)}\right), \\ H(A|B) &= H(A,B) - H(B) = \sum_l p(b_l)H(A|b_l) \\ &= \sum_l p(b_l)\sum_k p(a_k|b_l)\log\left(\frac{1}{p(a_k|b_l)}\right) \end{aligned} \tag{2.3}$$

Here, the entropy $H(A)$ corresponds to the prior uncertainty and the conditional entropy $H(A|B)$ corresponds to the amount of uncertainty remaining after observing features from sensor 2. The mutual information, that is defined as $I(A, B) = H(A) - H(A|B)$, corresponds to uncertainty that is resolved by observing B; in other words, features from sensor 2. From the definition of mutual information, it can be seen that the uncertainty that is resolved basically depends on the conditional entropy. Let us consider two types of sensors. Let the set of features of these two sensors be B_1 and B_2, respectively. If $H(A|B_1) < H(A|B_2)$ then $I(A, B_1) > I(A, B_2)$. This implies that the uncertainty is better resolved by observing B_1 as compared to B_2. This further implies that B_1 corresponds to features from a good sensor that is consistent with the features from sensor 1 and thus helps in improving the decision accuracy of sensor 1. B_2 corresponds to features from a bad sensor that is inconsistent with sensor 1, and hence, B_2 should not be considered.

Note that even though in the above example only two sensor nodes are considered for simplicity, this measure or metric can be used in a network of more than two sensors.

2.2. Euclidean Distance

Unlike mutual information, Euclidean distance does not evaluate the amount of information available from a second source. It does, however, measure the similarity between two feature sets in Euclidean space. This value can then be used to determine when to fuse two sources, whether from the same node or different nodes. A simple measure, Euclidean distance is defined as:

$$d = \sqrt{\sum_i (a_i - b_i)^2} \tag{2.4}$$

where a_i, b_i, and i are defined in Section 3.1.1.

2.3. Correlation

Correlation is also a well known measure of similarity. We use the standard measure of correlation as defined by:

$$\rho = \frac{E[(a-\mu_a)(b-\mu_b)]}{E[a-\mu_a]E[b-\mu_b]} \tag{2.5}$$

where μ_a and μ_b are the means of feature sets a and b, respectively. Note that correlation is very closely related to mutual information, $I(x,y)$, because (3.2) can be rewritten as:

$$I(x,y) = \sum_k p(a_k, b_k) \log\left(\frac{p(a_k, b_k)}{p(a_k)p(b_k)}\right). \tag{2.6}$$

2.4. Relative Entropy – Kullback-Leibler Distance

Finally, the Kullback-Liebler (KL) distance is derived from entropy, and again is a measure of the separation of two feature sets. It is defined as:

$$D = \sum_k p(a_k) \log\left(\frac{p(a_k)}{p(b_k)}\right) + \sum_k p(b_k) \log\left(\frac{p(b_k)}{p(a_k)}\right). \tag{2.7}$$

For example, to verify the discriminative power of different features that are used in the classifier, both acoustic and seismic data from SITEX02 was used for three vehicles – AAV, Dragon Wagon (DW) and HMMWV. The KL distance described above was computed for AAV versus DW, AAV versus HMMWV, and DW versus HMMWV. In Figure 2.1, KL distance versus acoustic and seismic features are plotted for these three cases. In this Figure, the first 20 on the X-axis correspond to acoustic features and the second 20 (21-40) correspond to seismic features. From this Figure, it can be seen that the first 10 features of acoustic and the first 10 features of seismic are most discriminative as compared to the last ten features, because they do not overlap. This indicates that they help in uniquely providing information between classes of targets.

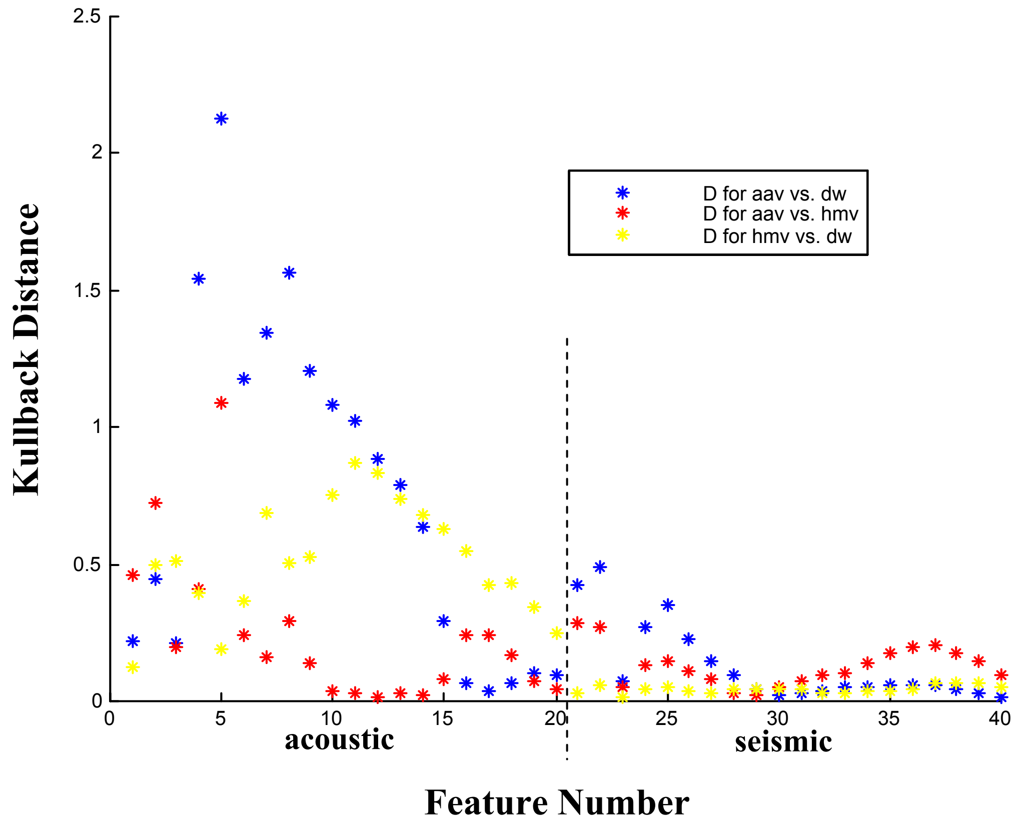

Figure 2.1 Discrimination power of individual features

3. Chapter: Review of Algorithms Used for Verification of Measures of Value of Information

The above described metrics are used to measure the value of information obtained from other sources such as multiple sensors on a single node and from the neighboring nodes in the context of target detection and classification. For target detection, an energy based detector is used and for classification, the maximum likelihood based classifier is used. As mentioned before, the value of information is in terms of improvement in the decision accuracy, which corresponds to classification accuracy for a classifier, and detection accuracy or probability of detection for a detector. Note that in this study, we did not develop a classifier or a detector. Instead, those developed by others in the SensIT program were used, since the goal of this study was to develop measures of value of information and to verify them by analyzing the detection and classification performances. In other words, the goal was to find the lower and upper bound on the performance of these algorithms. In the following two sections, we review the classifier and the detector that were used in this study.

3.1. Maximum likelihood based classifier:

The classifier we used for the verification of measures of value of information in terms of improving the decision accuracy is a maximum likelihood based classifier developed by the University of Wisconsin [4] as part of this SensIT program. For a given set of training features and target labels, a Gaussian mixture model was determined during the training phase of the classifier. During testing, the distance between the test feature vector and the i^{th} class Gaussian mixture is computed. This corresponds to negative log likelihood. Then, *a priori* probability is used to obtain the maximum a posterior classification. The features set that was used here consists of twenty features from the power spectral density. This was computed using a 1024 FFT. The feature set was collected by summing up the values over equal length segments of the power spectrum. For the acoustic and seismic sensors, the maximum frequency used was 1000 and 200 Hz, respectively.

3.2. Energy based detector:

An energy based detector was also used for the verification of improvement in decision accuracy when the value of information based fusion architecture was used. This detector was developed by BAE, Austin [5]; also as part of the SensIT program. A brief description of this detector was provided below.

For every block of a given signal, the energy of the down sampled version of the power spectral density was computed. For the computation of the power spectral density, a 1024 point FFT was used. This energy was compared with a threshold value. Whenever the energy was above the threshold, it was declared that the target was detected. The threshold value was adaptively changed based on the background energy.

4. Chapter: Experimental details

The above described classifier and detector, measures of value of information, and the fusion algorithm, which uses these measures while deciding when to and when not to fuse information, were implemented in Matlab; a product of MathWorks. They were tested using real data that was collected by distributing sensor nodes along the east-west and north-south road at Twentynine Palms, CA, during one of the field tests (SITEX'02). These sensor nodes were manufactured by Sensoria. On each sensor node, three sensors - acoustic, seismic and IR, a four channel data acquisition board, and a processing board are available. These nodes also have communication capabilities. For more details on the sensor node, refer to [6].

SITEX'02 data corresponds to acoustic, seismic, and IR data of three vehicles – AAV, Dragon Wagon (DW) and HMMWV, moving along the east-west and north-south road as shown in Figure 4.1. In this figure, nodes placements are also provided. In all twenty-four nodes out of sixty-nine nodes were considered in our experiments, because the data from the nodes that were not considered was not good; i.e., the sensors on these nodes did not operate as desired. Both seismic and acoustic data was used, from the nodes that were considered.

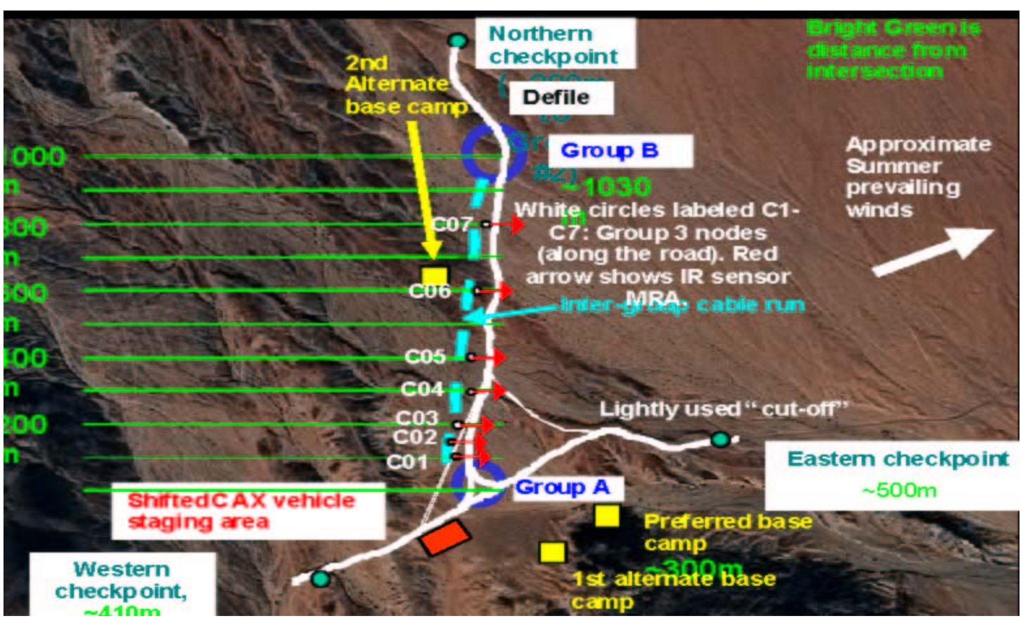

Figure 4.1 Distribution of Sensor nodes at Twenty nine Palms, CA

In the next section, the classification experimental details and the results are provided. In section 4.2, the detection experiments and the results are provided. In both of these sections the experimental details and results are provided with and without the value of information based fusion technique that was developed in this study.

4.1. Classification Experiments

First, acoustic data from each node is considered. The maximum likelihood classifier is trained using only acoustic data from individual nodes. The challenges in the classification experiments are threefold: 1) when to reject a source of data, 2) when to propagate data between sequential nodes, and 3) when to share individual sensor data within the same node. Using only acoustic data, we investigated the effectiveness of the four measures of value of information outlined in Chapter 2 - mutual information, Euclidean distance, correlation, and Kullback-Liebler distance.

In addition, we investigated two methods of using these measures. When evaluating the effectiveness of fusing two sources of data, is it better to compare the two sources with each other or with the stored training data? To answer this question, we devised several similarity measures to measure the closeness of two data sources. We calculated these measures between data at all sequential nodes. Then, for each similarity measure, we computed its correlation with correct classification performance at each node. We call this the performance correlation. The average performance correlation over all the nodes for each class of data, using previous node similarity measures, is shown in Figures 4.2 and 4.3 for acoustic and seismic sensors, respectively. In these two figures, the numbering on the X-axis is artificial in that 1 corresponds to AAV, 2 corresponds to Dragon wagon, and 3 corresponds to HMMWV.

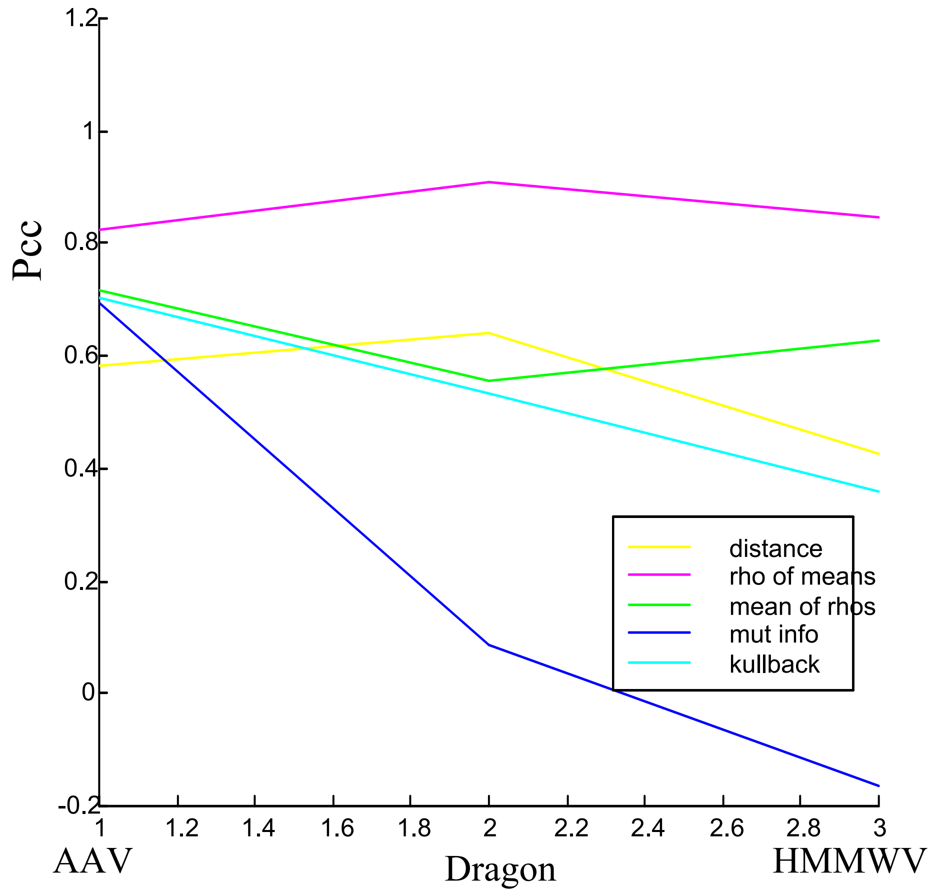

Figure 4.2: Performance (Pcc – classification probability) correlation with previous node for acoustic sensor

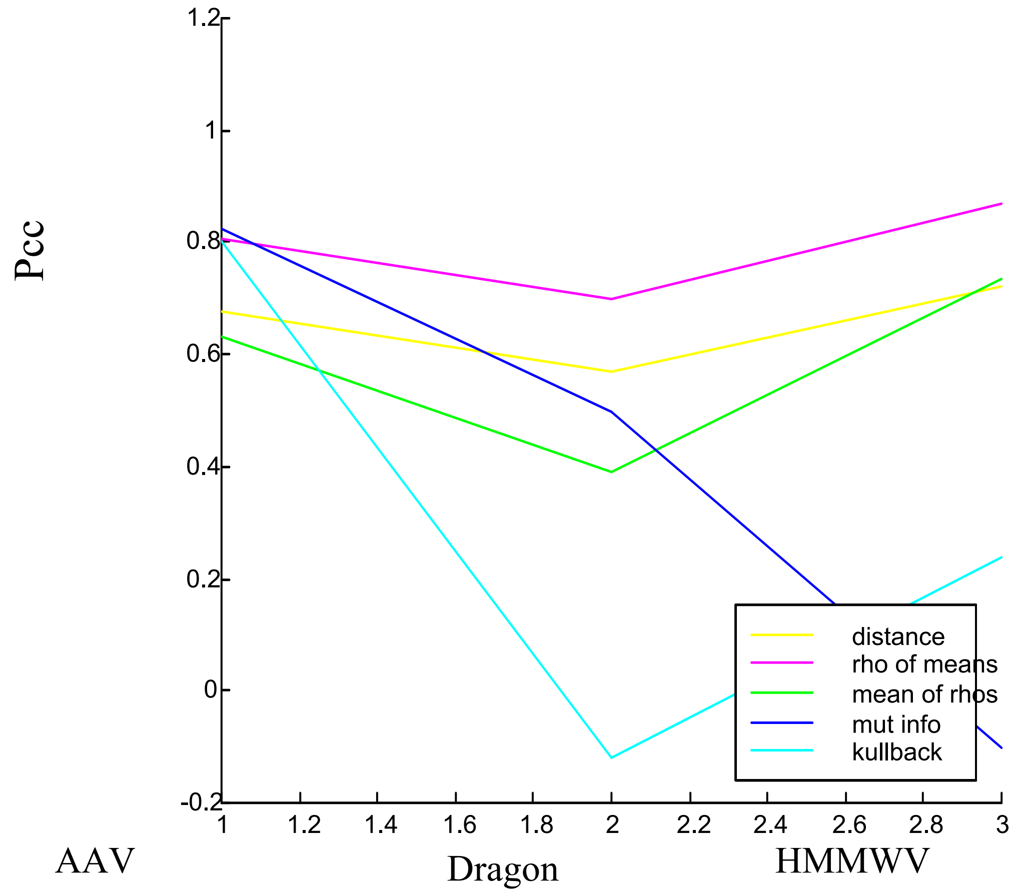

Figure 4.3: Performance (Pcc – classification probability) correlation with previous node for seismic sensor

Next, we calculated the same similarity measures between the data at each node and the data stored in the training sets. Again, for each similarity measure, we computed its correlation with correct classification performance at each node.

The average performance correlation over all nodes for each class of data, using training set similarity measures, is shown in Figure 4.4 for acoustic sensors. Inspection of Figures 4.2, 4.3 and 4.4 shows that the similarity measures Euclidean distance and correlation are more closely aligned with the correct classification performance, than either mutual information or Kullback-Liebler distance. In practice, however, we found that the Euclidean distance outperformed correlation as the determining factor in fusion decisions.

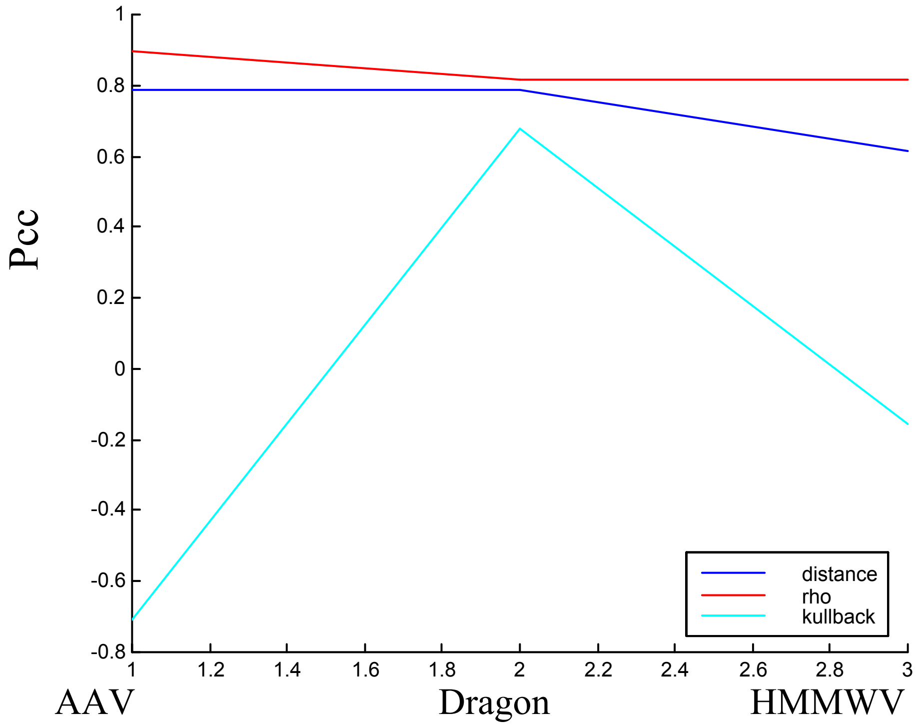

Figure 4.4: Performance (Pcc – classification probability) correlation with training class data for acoustic sensor

Furthermore, comparing Figures 4.2 and 4.4, shows that using the training set for similarity measures is more effective than using the data from the previous node in the network. We found this to be true in practice as well. Subsequent work with the seismic data echoed the findings of the acoustic data as evidenced by Figure 4.3. Note that even though we use the training data to make the fusion decision, we perform the actual data fusion with current and previous node data.

4.1.1. Rejection of bad data:

Sometimes, one node or one sensor can have bad data, in which case we prefer to reject this data rather than classify with poor results. We investigated one way of recognizing such sources of bad data, by observing outliers in each dimension of the 20-dimensional feature vectors. First, we computed the mean of the data at the node in question. The value at each dimension was then compared to the mean for that dimension of the stored training sets. If the data contained an outlier in 4 or more of the dimensions, for each of the training classes, we rejected the data. By rejecting the data, we did not fuse it with any other data, pass it on to any other node, nor even compute a classification at that source. Our method resulted in the rejection of several sources of bad data, thus improving the overall classification results as shown in Figures 4.5 and 4.6. In these figures, the X-axis corresponds to the node number and the Y-axis corresponds to classification accuracy. Further, the blue plot corresponds to classification accuracy without fusing information from neighboring nodes, whereas the red plot corresponds to classification accuracy by fusing information, using the measures of value of information described before. From these figures, it can also be seen that by fusing information based on its value, the classification accuracy reaches close to 100 % asymptotically.

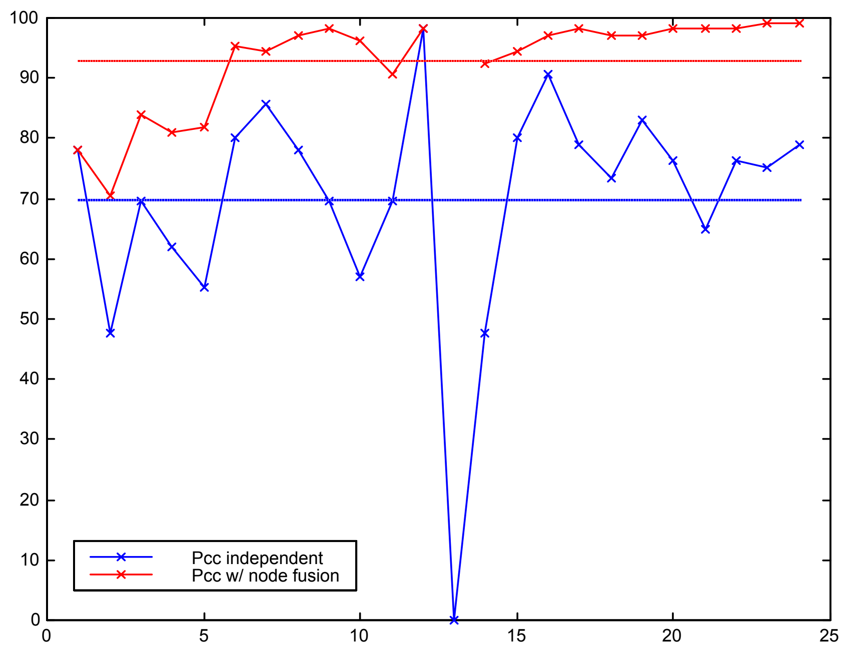

Figure 4.5: Performance of node fusion for the AAV with only acoustic sensor data

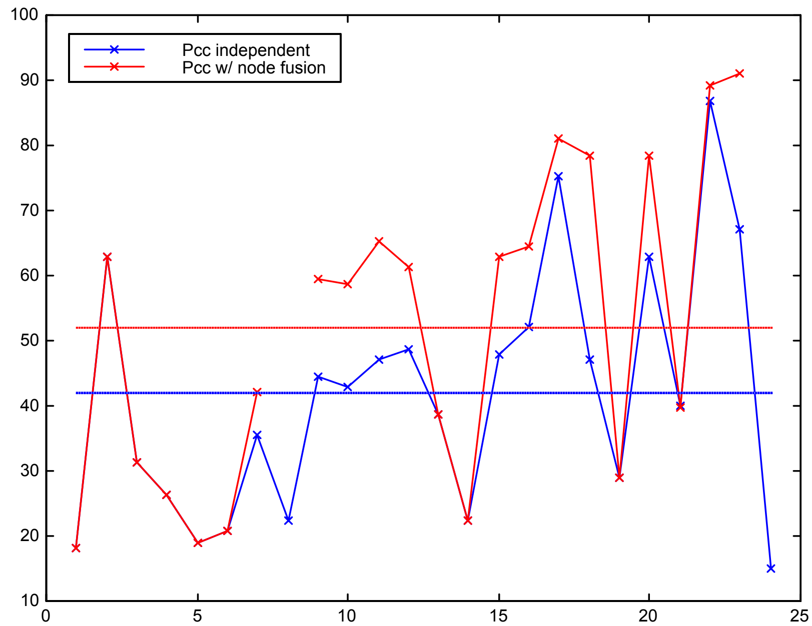

Figure 4.6: Performance of node fusion for the AAV with only seismic sensor data

4.1.2. Node to node fusion:

The fusion decision can be made with a threshold; i.e., if the distance between two feature sets is below some value, then fuse the two feature sets. The threshold value can be predetermined off-line or adaptively. We sidesteped the threshold issue, however, by basing the fusion decision on relative distances. To do so, we initially assume the current node belonged to the same class (aka the target class) as the previous node and employ the following definitions. Let x_n be the mean vector of the current node data. Let x_{nf} be the mean vector of the fused data at the current node. Let x_{c1} be the mean vector of the target training class data. Let x_{c2}, x_{c3} be the mean vectors of the remaining training classes. A Euclidean distance ratio is defined as:

$$r_{dist} = d_{c1} / \min(d_{c2}, d_{c3}), \tag{4.1}$$

where d_{ci} is the Euclidean distance (3.4) between x_n and x_{ci}. We then use the following pseudo-code to make our fusion decisions.

if (r_{dist} <= 1.0)

fuse_4class = 1;   fuse_4carry = 1;

class_ind = classify x_n;

```
        if (class_ind >= 70%)  check class_fuse;
        end
else
        fuse_4class = 0;    fuse_4carry = 0;
        if {(d_c1 <= 3s_c1 ) & (d_c2 <= 3s_c2 ) & (d_c2 <= 3s_c2 )}
                class_ind = classify x_n;
                if (class_ind == target class)  fuse_4class = 1;
                        if (class_ind >= 70%)
                                fuse_4carry = 1;
                                class_fuse = classify x_nf;
if (class_ind > class_fuse)
                                        class_fuse = class_ind;
end
                        end
                end
        else
                reject this data;
        end
end
```

There are two outcomes to the fusion decision. First, we decide whether or not to fuse the data at the current node. If the current node has bad data, fusion can pull up the performance. We may not want to carry the bad data forward to the next node (the second fusion decision outcome). *fuse_4class* is a flag indicating whether or not to fuse for the current classification. *fuse_4carry* is a flag indicating whether or not to include data from the current node in the fused data that is carried forward. In Figures 4.5 and 4.6, we show the correct classification improvement gained by fusing from node to node for the acoustic and seismic sensors, respectively. For the acoustic sensor, we show the classification results from the AAV data, while using the DW data for the seismic sensor results. In the case of the acoustic data, the mean correct classification performance across all nodes increases from 70%, for independent operation, to 93%, with node to node fusion across the network. Similarly, the seismic correct classification performance increases from 42% to 52%. However, the important thing to notice is the asymptotic increase in the classification accuracy from node to node, when the fusion is performed based on our measures of value of information.

4.1.3. Fusion between sensors:

After fusion from node to node of the individual sensors, we look at the benefit of fusing the acoustic and seismic sensor data at the same node. To do so, we employ the following definitions. Let r_{dist} be defined as in equation (4.1), but with the new data types (a - acoustic, s - seismic, and as – a concatenated acoustic/seismic vector). Let x_a be the mean vector of the current node acoustic data, after fusion from node to node. Let x_s be the mean vector of the current node seismic data, after fusion from node to node. Let $x_{as} = x_a$, concatenated with x_s (dumb fusion). Let x_{asf} = smart fusion of x_a with x_s. Let x_{in} be the data input to the classifier. Now, we employ two steps in the sensor fusion process as shown in the pseudocode below.

First we employ a smart sensor fusion routine:

indx = min(r_a_{dist}, r_s_{dist}, r_as_{dist})

if (indx == 1) $x_{in} = x_a$;

elseif (indx == 2) $x_{in} = x_s$;

elseif (indx == 3) $x_{in} = x_{as}$;

end

Next, we employ a final fusion routine:

class_acst = classify x_a;

class_seis = classify x_s;

class_as_dumb = classify x_{as};

class_as_smart = classify x_{asf};

if { (class_acst >= 70%) | (class_seis >= 70%) | (class_as_ind >= 70%) }

    class_final_fuse = max (class_acst, class_seis,   class_as_dumb, class_as_smart)

end

The classification performance is averaged over all the nodes for each vehicle class. The correct classification performance improves at each stage of fusion processing as shown in Table 4.1.

The results indicate that the fusion based on value of information helps in improving the decision accuracy at each node significantly.

	AAV	DW	HMMV
Acoustic independent	70 %	58 %	46%
Seismic independent	72%	42 %	24%
Acoustic fusion	93%	80%	69%
Seismic fusion	93%	52%	31%
Acoustic & seismic, Independent	76%	55%	58%
Acoustic & seismic, With fusion	95%	90%	77 %

Table 4.1: Summary of classification performance

Furthermore, we studied the implication of combination of fusion, with and without using measures of value of information and rejection capability in a systematic way, for each sensor type and for both sensors, by using four modes of operation which are described below.

- Mode 1 corresponds to each node or sensor operating independently of the others (i.e., no fusion), without rejection capability.
- Mode 2 corresponds to each node (single sensor on each node or multiple sensors on each node) propagating information from one node to the other; its usage dependent on our value of information measures, without rejection capability.
- Mode 3 corresponds to each node operating independently (i.e., no fusion), with rejection capability.
- Mode 4 corresponds to each node (single sensor or multiple sensors) propagating information from one node to the other; its usage based on our value of information measures, with rejection capability.

In the following tables, the correct classification performance for each vehicle – AAV, DW, and HMMWV is listed for each of the modes described above. The overall average classification performance that we can obtain by using both, value of information measures while fusing information and by rejecting bad data, is shown in Figure 4.7. From this Figure, it can be seen that the classification accuracy improves significantly and we can reach the upper bound of close to 100 % classification accuracy.

Acoustic Sensor Alone

	AAV	DW	
Mode 1	70	58	46
Mode 2	90	73	69
Mode 3	74	76	49
Mode 4	95	96	77

Table 4.2: The classification performance under four different operations using only acoustic sensor

Seismic Sensor Alone

	AAV	DW	
Mode	72	42	24
1	89	44	33
Mode	81	61	30
2	95	66	44

Table 4.3: The classification performance under four different operations using only seismic sensor

Acoustic and Seismic Blind Fusion

	AAV	DW	
Mode 1	76	55	58
Mode 2	93	64	71
Mode 3	80	74	42
Mode 4	97	87	63

Table 4.4: The classification performance under four different operations by blindly fusing information (i.e., without using our value of information measures) between acoustic & seismic sensors, and from neighboring nodes

Acoustic and Seismic Smart Fusion

	AAV	DW	
Mode 1	80	69	62
Mode 2	94	86	85
Mode 3	84	81	52
Mode 4	98	95	83

Table 4.5: The classification performance under four different operations by fusing information smartly (i.e., using our value of information measures) between acoustic & seismic sensors, and from neighboring nodes

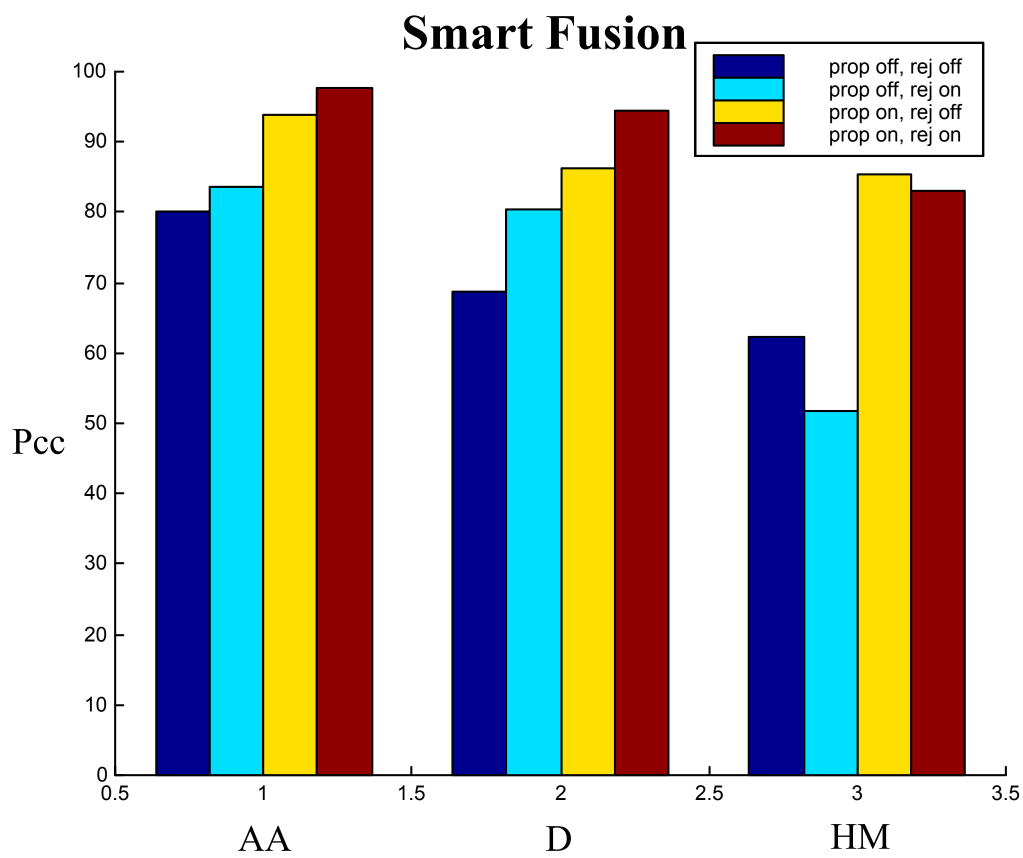

Figure 4.7: The overall average classification accuracy for three vehicles under four modes using both measures of value of information and rejection of bad data while fusing information between sensors and from the neighboring nodes

In the above figure, "prop" means propagation of information from one node to the other and "rej on/off" means rejection algorithm used or not used.

In short, the overall observations of the analysis of the maximum likelihood based classifier, using our measures of value of information and our strategy for rejection of bad data, is as follows:

1. Propagation always helps.
2. For independent sensors, rejection always helps.
3. Smart fusion always improves on blind fusion.
4. Blind fusion is not always better than acoustic alone.

5. Smart fusion is always better than acoustic alone and blind fusion.
6. Rejection is always better when using sensor fusion, except in the case of the HMMWV data.

4.2. Detection experiments:

For the detection experiments, both acoustic and seismic data were again considered. First, only acoustic data from individual nodes were used. A threshold value was initially set, which was varied adaptively based on the background energy. The power spectral density of acoustic data was computed using a 1024 point FFT and it was downsampled by 8. The energy of the downsampled version of the power spectral density was computed. This energy was compared with the threshold value. If the energy was above the threshold, it was decided that the target was detected. The time of detection and the confidence on detection were also calculated. The detection and time of detection were compared with the ground truth. If the target was detected when it was supposed to be, and if the time of detection was within the region of interest, then it was counted towards calculating the probability of detection. If the detection time was outside the region of interest (missed detection), and if a target was detected when it should not have been (false alarm), they were counted towards computing the probability of false alarm. The probabilities of detection and false alarm, using only acoustic data from individual nodes, without any fusion for AAV, DW and HMMWV were: 0.8824, 0.8677, 0.8382 & 0.1176, 0.1323, 0.1618, respectively. Similarly, the probabilities of detection and false alarm, using only seismic data from individual nodes, without any fusion for AAV, DW and HMMWV were: 0.8030, 0.7910, 0.5735 & 0.1970, 0.2090, 0.4265, respectively.

Next, the mutual information based value of information measure was used on the energy of power spectral density to make a decision of fusing data between sensors - acoustic and seismic on each individual node. The detector was tested using the fused data on each node. The probability of detection and false alarm were computed as described above. The probabilities of detection of this intelligently fused data for AAV, DW and HMMWV were: 0.9394, 0.9105 and 0.8529, respectively. The probabilities of false alarm is not provided here because it is equal to 1 – probability of detection, since both false alarm and missed detections are combined together.

These results are summarized in Figure 4.8 in the form of a bar graph. From this, it can be seen that the intelligent sensor data fusion, based on value of information, significantly improves the detection accuracy. This type of fusion especially helps in difficult data as in the case of HMMWV.

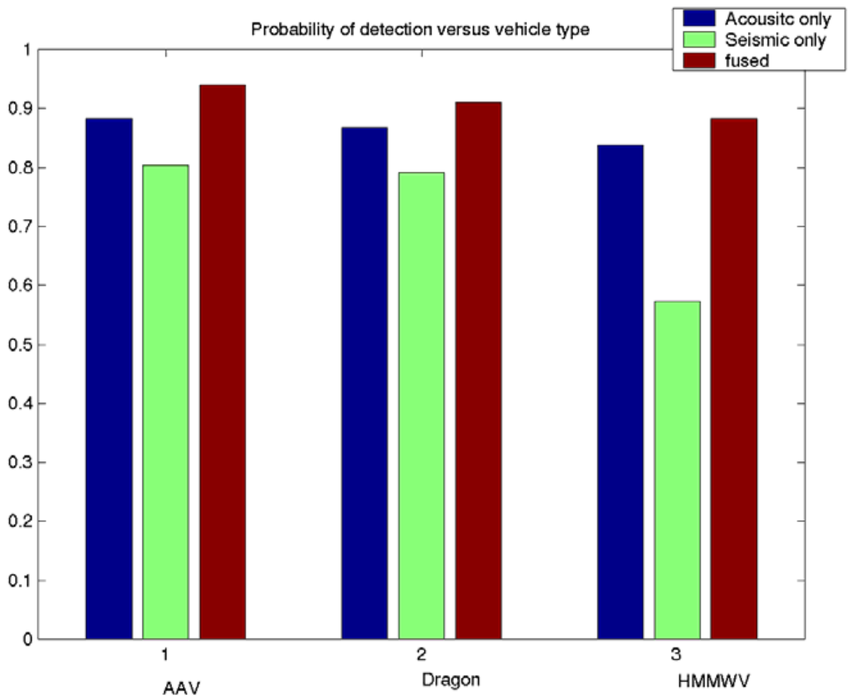

Figure 4.8: The performance of an energy based detector

Detailed analysis similar to the classifier described above, lead to similar observations in the case of a detector. These two analyses, using real data, imply that the sensor/data fusion using value of information measures and rejection strategies significantly improve the decision accuracy. This would have a great impact on DoD applications, such as automatic target detection and recognition, surveillance, and situation awareness. In addition to improving the decision accuracy, this study provides a technique to efficiently manage and monitor sensors in a distributed network. This would have an impact in generating a single integrated picture that could be used in situation awareness applications.

5. Chapter: Summary and future directions

5.1. Summary:

In this final report, the work done by HRL Laboratories, LLC under the contract F30602-01-C-0192 has been included. In short, this work has resulted in measures for value of information and application of these measures in making decisions on when to and when not to fuse information between different sensors types and information from the neighboring nodes, and when to reject bad data in a network of distributed sensors. In particular, the significant achievements of this work were as follows:

- Measures based on Euclidean distance, Correlation, Mutual information and relative entropy (Kullback-Liebler) were developed (Chapter 2).
- An energy based detector and a maximum likelihood based classifier were considered (Chapter 3).
- The above mentioned measures were used in assessing the value of information obtained from multiple sensors on the same node and from the neighboring nodes. The value of information was in terms of improvement in decision accuracy, which corresponds to detection or classification accuracy in the context of a detector or a classifier. If the information obtained from the other sensors of the same node or from the neighboring nodes did not improve the decision accuracy then it was discounted. Otherwise, it was fused with the existing information (Chapter 4).
- The measures and the performance of the energy based detector and the maximum likelihood based classifier were evaluated using the SITEX02 data (Chapter 4).
- The performance analysis included using (a) only one sensor on each node with and without fusing information from the neighboring nodes, based on whether the value is added or not and with and without rejection capability; the performance without fusion corresponds to the lower bound, (b) multiple sensors on each node with and without fusing information from the other sensors, based on whether the value is added or not and with and without bad data rejection capability, and (c) multiple sensors on each node and the neighboring nodes with and without fusion and with and without the capability

to reject the bad data; the performance with fusion of information from multiple sensors and the neighboring nodes when values is added corresponds to the upper bound.

- The experimental results indicated that by fusing information from multiple sensors on the same node and from the neighboring nodes when the value is added, the average performance of both a detector and a classifier improved significantly and the upper bound of close to 100 % decision accuracy can be obtained (Chapter 4).

- An invention disclosure based on this study has been submitted. Also a technical paper was published in the proceedings of the Fusion'03 conference, held in Cairns, Australia from July 8-11, 2003.

5.2. Future work:

The measures of value of information developed under this study can be used for fusion for different types of sensors. We have submitted a proposal in response to DARPA/DSO's BAA on time-reversal methods, in which we have proposed to use the measures of value of information in fusing images obtained from different looks of distributed multi-static millimeter wave radar sensors to generate a 3D map or image for an enhanced see through the wall surveillance application. If we win this contract, we will further develop these measures. We have also been using these measures in a company funded project to develop techniques for distributed fusion. In this project, visualization and network aspects are also addressed. The final goal of this project is to generate an enhanced single integrated picture for a situational awareness application. We will explore applying value of information measures to develop higher level fusion such as perception and cognition level fusion techniques. We will also explore the application of these measures or related measures for sensor or resource selection and management under this project. We believe that this DARPA contract has resulted in the powerful concept of "value of information", which could be used in many different applications and also has resulted in an analysis strategy which could be used to assess the performance of several decision making algorithms.

6. References

1. S. Kadambe, ``Information theoretic based sensor discrimination for information fusion and cluster formation in a network of distributed sensors," in *Proc. of 4th annual conference on information fusion*, Montreal, Quebec, Canada, August 7-10, 2001, pp. ThC1-19-ThC1-25.

2. S. Kadambe, ``Feature discovery and sensor discrimination in a network of distributed sensors for target tracking," in *Proc. of IEEE workshop on Statistical signal processing*, Singapore, August 6-8, 2001, pp. 126-129.

3. R. Battti, "Using mutual information for selecting features in supervised neural net learning," *IEEE Trans. On Neural Network*, vol. 5, no. 4, July 1994, pp. 537-550.

4. S. C. A. Thomopoulos, "Sensor selectivity and intelligent data fusion," *Proc. Of the IEEE MIT'94*, October 2-5, 1994, Las Vegas, NV, pp. 529-537.

5. J. Manyika and H. Durrant-Whyte, Data fusion and sensor management: An information theoretic approach, Prentice Hall, 1994.

6. Papoulis, Probability, Random variables and Stochastic Processes, Second edition, McGraw Hill, 1984, pp. 500-567.

7. Y. Hen Wu, "Maximum likelihood based classifier," *SensIT PI meeting,* Jan 15-17, 2002, Santa Fe, New Mexico.

8. S. Beck, "Energy based detector," *SensIT PI meeting,* Jan 15-17, 2002, Santa Fe, New Mexico.

9. www.sensoria.com

10. S. Kadambe and C. Daniell, "Value of information based sensor/data fusion," in *proc. Of 6th annual conference on information fusion*, Cairns, Australia, Jul 8-11, 2003.

Sensor/Data Fusion Based on Value of Information [1]

Shubha Kadambe and Cindy Daniell

HRL Laboratories, LLC
3011 Malibu Canyon Road, Malibu CA 90265, USA
E-mail: {skadambe,daniell}@hrl.com

Abstract - *Spatially distributed network of inexpensive, small and smart nodes with multiple onboard sensors is an important class of emerging networked systems for various applications. Since this network of sensors has to operate efficiently in adverse environments, it is important that these sensors process information efficiently and share information such that the decision accuracy is improved. One way to address this problem is to measure the value of information obtained from multiple sensors on the same sensor node as well as from the neighboring nodes and fuse that information if value is added in terms of improvement in decision accuracy. In this paper, the measures that are developed for assessing the value of information are described. These measures are then used in making a decision of fusing either features or data from multiple sensors and neighboring nodes. While making this decision whether the value is added by fusing the information, is verified by conditioning it on improving the decision accuracy. The measures and value added are verified by using real data collected at Twentynine Plams, CA, USA and in the context of target detection and classification. From the results of improvement in classification accuracy and probability of detection reported in this paper, it can be seen that the utilization of measure of value of information while fusing helps in improving the decision accuracy significantly.*

Keywords: Value of information, measures of value, mutual information, decision accuracy, sensor/data fusion.

1 Introduction

Spatially distributed network of inexpensive, small and smart nodes with multiple onboard sensors is an important class of emerging networked systems for various defense and commercial applications. Since this network of sensors has to operate efficiently in adverse environments using limited battery power and resources, it is important that these sensors process information efficiently and share information such that the decision accuracy is improved. In this paper, this is addressed by developing measures that assess the value of information obtained from multiple sensors on board on a node and from the neighboring nodes by conditioning it on improvement in the decision accuracy. If the information obtained from other sensor types on a node and/or from the neighboring nodes do improve the decision accuracy then the information is fused. In our study, information is obtained in the form of features (for classification) or data (for detection). In [1-2] we have developed a general information theoretic based metric that can be used in any kind of sensor selection and data fusion. However, while analyzing the real data with respect to a classifier and a detector we observed that the correlation between the metric of value of information and the decision accuracy depend on the type of a classifier or a detector. Hence, we think that mutual information metric may not work always. Therefore, in this paper, we have developed several measures and studied them systematically with respect to one type of classifier and a detector.

To our knowledge this value of information based fusion is not studied by others and is the significant contribution of this paper. In [3], the author shows that in general by fusing data from selective sensors the performance of a network of sensors can be improved. However, this study does not describe specific novel measures for value of information and data fusion based on the assessed value unlike in this paper. In [4], techniques to represent Kalman filter state estimates in the form of information – Fisher and Shannon entropy are provided. In such a representation it is straightforward to separate out what is new information from what is either prior knowledge or common information. This separation procedure is used in decentralized data fusion algorithms that are described in [5]. This is different from our paper in that we have developed measures for value of information and perform sensor/data fusion if the added value is in terms of improving the decision accuracy. The rest of the paper is organized as follows: In the next section details of measures that we have developed are described. Section 3 provides a brief description of the classifier and the detector that we have used in our study for the purposes of verification of the measures and data/sensor fusion. Section 4 provides the details of real data that we use in this study, the experimental setup and results. In section 5, we conclude and provide future research directions.

2 Measures of value of information

This section describes the measures of value of information that we have developed. Even though the mathematics of the metrics described below are not novel, the usage of metrics in the context of verifying value of information with respect to improving the decision accuracy (e.g., classification accuracy, detection accuracy) is new.

2.1 Mutual information:

Entropy is a measure of uncertainty. Let $H(x)$ be the entropy of previously observed x events. Let y be a new event. We can measure the uncertainty of x after including y by using the conditional entropy which is defined as:

$$H(x|y) = H(x, y) - H(y) \qquad (1)$$

with the property $0 \leq H(x|y) \leq H(x)$. The conditional entropy $H(x|y)$ represents the amount of uncertainty remaining about x after y has been observed. If the uncertainty is reduced then there is information gained by observing y. Therefore, we can measure the value of y by using conditional entropy. Another measure that is related to conditional entropy that one can use is the mutual information $I(x,y)$ which is a measure of uncertainty that is resolved by observing y and is defined as:

$$I(x,y) = H(x) - H(x|y). \qquad (2)$$

To explain how this measure can be used to measure value of information obtained from another sensor type an example is provided below.

2.1.1 An example of value of information using mutual information as a metric

Let $A = \{a_k\}$ k = 1, 2,… be the set of features from sensor 1 and let $B = \{b_l\}$ l = 1, 2,… be the set of features from sensor 2 on the same node. Let $p(a_i)$ be the probability of feature a_i. Let $H(A)$, $H(B)$ and $H(A|B)$ be the entropy corresponding to sensor 1, sensor 2 and sensor 1 given sensor 2, respectively, and they are defined as [6]:

$$H(A) = \sum_k p(a_k)\log\left(\frac{1}{p(a_k)}\right),$$
$$H(A|B) = H(A,B) - H(B) = \sum_l p(b_l)H(A|b_l) \qquad (3)$$
$$= \sum_l p(b_l)\sum_k p(a_k|b_l)\log\left(\frac{1}{p(a_k|b_l)}\right)$$

Here, $H(A)$ the entropy corresponds to the prior uncertainty and $H(A|B)$ the conditional entropy corresponds to the amount of uncertainty remaining after observing features from sensor 2. The mutual information that is defined as $I(A, B) = H(A) - H(A|B)$ corresponds to uncertainty that is resolved by observing B in other words features from sensor 2. From the definition of mutual information, it can be seen that the uncertainty that is resolved basically depends on the conditional entropy. Let us consider two types of sensors. Let the set of features of these two sensors be B_1 and B_2, respectively. If $H(A|B_1) < H(A|B_2)$ then $I(A, B_1) > I(A, B_2)$. This implies that the uncertainty is better resolved by observing B_1 as compared to B_2. This further implies that B_1 corresponds to features from a good sensor that is consistent with the features from sensor 1 and thus helps in improving the decision accuracy of sensor 1 and B_2 corresponds to features from a bad sensor that is inconsistent with sensor 1 and hence, B_2 should not be considered.

Note that even though in the above example only two sensor nodes are considered for simplicity, this measure or metric can be used in a network of more than two sensors.

2.2 Euclidean Distance

Unlike mutual information, Euclidean distance does not evaluate the amount of information available from a second source. It does, however, measure the similarity between two feature sets in Euclidean space. This value can then be used to determine when to fuse two sources, whether from the same node or different nodes. A simple measure, Euclidean distance is defined as:

$$d = \sqrt{\sum_i (a_i - b_i)^2} \qquad (4)$$

where a_i, b_i, and i are defined in Section 2.1.1.

2.3 Correlation

Correlation is also a well known measure of similarity. We use the standard measure of correlation as defined by:

$$r = \frac{E[(a - m_a)(b - m_b)]}{E[a - m_a]E[b - m_b]} \qquad (5)$$

where m_a and m_b are the means of feature sets a and b, respectively. Note that correlation is very closely related to mutual information, $I(x,y)$ because (2) can be rewritten as:

$$I(x,y) = \sum_k p(a_k, b_k) \log\left(\frac{p(a_k, b_k)}{p(a_k)p(b_k)}\right). \qquad (6)$$

2.4 Kullback-Liebler Distance

Finally, the Kullback-Liebler (KL) distance is derived from entropy, and again is a measure of the separation of two feature sets. It is defined as:

$$D = \sum_k p(a_k) \log\left(\frac{p(a_k)}{p(b_k)}\right) + \sum_k p(b_k) \log\left(\frac{p(b_k)}{p(a_k)}\right). \qquad (7)$$

3 Review of Algorithms used for verification

The above described metrics are used to measure the value of information obtained from other sources such as multiple sensors on a single node and from the neighboring nodes in the context of target detection and classification. For target detection, energy based detector was used and for classification, maximum likelihood based classifier was used. As mentioned before the value of information is in terms of improvement in the decision accuracy which corresponds to classification accuracy for a classifier and detection accuracy or probability of detection for a detector. Note that in this study, we did not develop a classifier or a detector; however, used those developed by others since the goal of this study is to develop measures of value of information and verify them in terms of improvement in decision accuracy when they were used to make a decision of whether to fuse information obtained from the other source or not. In the following two sections we review the classifier and the detector that were used in this study.

3.1 Maximum likelihood based classifier

The classifier we used for the verification of measures of value of information in terms of improving the decision accuracy is a maximum likelihood based classifier developed by the University of Wisconsin [7] as part of DARPA's sensor information technology (SensIT) program. For a given training features and target labels a Gaussian mixture model is determined during the training phase of the classifier. During testing the distance between the test feature vector and ith class Gaussian mixture is computed. This corresponds to negative log likelihood. Then *a priori* probability is used to obtain the maximum a posterior classification. The features' set that is used here consists of twenty features from the power spectral density. This is computed using 1024 FFT. The feature set is collected by summing up the values over equal length segments of the power spectrum. For the acoustic and seismic sensors the maximum frequency used was 1000 and 200 Hz, respectively.

3.2 Energy based detector

An energy based detector is also used for the verification of improvement in decision accuracy when the value of information based fusion architecture is used. This detector is developed by BAE, Austin [8]; also as part of the SensIT program. A brief description of this detector is provided below.

For every block of a given signal the energy of the down sampled version of the power spectral density is computed. For the computation of the power spectral density, 1024 point FFT is used. This energy is compared with a threshold value. Whenever the energy is above the threshold it was declared that the target was detected. The threshold value is adaptively changed based on the background energy.

4 Experimental details

The above described classifier and detector, and measures of value of information and the fusion algorithm which uses these measures while deciding when to and when not to fuse information were implemented in Matlab, a product of MathWorks and were tested using real data that was collected by distributing sensor nodes along east-west and south-north road at Twentynine Palms, CA during one of the field tests (SITEX'02). These sensor nodes are manufactured by Sensoria. On each sensor node, three sensors - acoustic, seismic and IR sensors, a four channel data acquisition board and a processing board are available. These nodes also have communication capabilities. For more details on the sensor node, refer to [9].

SITEX'02 data corresponds to acoustic, seismic and IR data of three vehicles – AAV, Dragon Wagon (DW) and HMMWV moving along the east-west and north-south road as shown in Figure 1. In this figure, nodes placements are also provided. Totally twenty four nodes were considered in our experiments. We used both seismic and acoustic data from these nodes. In the next section, the classification experimental details and the results are provided and in section 4.2 the detection experiments and the results are provided. In both these sections experimental details and results are provided with and without value of information based fusion technique that was developed in this study.

4.1 Classification experiments

First, acoustic data from each node is considered. The maximum likelihood classifier is trained using only acoustic data from individual nodes. The challenges in the classification experiments are threefold: 1) when to reject a source of data, 2) when to propagate data between sequential nodes, and 3) when to share individual sensor data within the same node. Using only acoustic data, we investigated the effectiveness of the four measures of value of information outlined in Section 2 - mutual information, Euclidean distance, correlation, and Kullback-Liebler distance.

In addition, we investigated two methods of using these measures. When evaluating the effectiveness of fusing two sources of data, is it better to compare the two sources with each other or with the stored training data? To answer this question, we devised several similarity measures to measure the closeness of two data sources. We calculated these measures between data at all sequential nodes. Then for each similarity measure, we computed its correlation with correct classification performance at each node. We call this the performance correlation. The average performance correlation over all nodes for each class of data using previous node similarity measures is shown in Figure 2. Next, we claculated the same similarity measures between the data at each node and the data stored in the training sets. Again, for each similarity measure, we computed its correlation with correct classification performance at each node.

The average performance correlation over all nodes for each class of data using training set similarity measures is shown in Figure 3. Inspection of Figures 2 and 3 show that the similarity measures Euclidean distance and correlation are more closely aligned with correct classification performance than either mutual information or Kullback-Liebler distance. In practice, however, we found that the Euclidean distance outperformed correlation as the determining factor in fusion decisions. Furthermore, comparing Figures 2 and 3 shows that using the training set for similarity measures is more effective than using the data from the previous node in the network. We found this to be true in practice as well. Subsequent work with the seismic data echoed the findings of the acoustic data. Note that even though we use the training data to make the fusion decision, we perform the actual data fusion with current and previous node data.

4.1.1 Rejection of bad data

Sometimes one node or one sensor can have bad data, in which case we prefer to reject this data rather than classify with poor results. We investigated one way of recognizing such sources of bad data, by observing outliers in each dimension of the 20-dimensional feature vectors. First we computed the mean of the data at the node in question. The value at each dimension was then compared to the mean for that dimension of the stored training sets. If the data contained an outlier in 4 or more of the dimensions, for each of the training classes, we rejected the data. By rejecting the data, we did not fuse it with any other data, pass it on to any other node, nor even compute a classification at that source. Our method resulted in the rejection of several sources of bad data, thus improving the overall classification results as shown in Figures 4 and 5.

4.1.2 Node to node fusion

The fusion decision can be made with a threshold, i.e. if the distance between two features sets is below some value, then fuse the two feature sets. The threshold value can be predetermined off-line or adaptive. We sidestep the threshold issue, however, by basing the fusion decision on relative distances. To do so, we initially assume the current node belonged to the same class (aka the target class) as the previous node and employ the following definitions. Let x_n be the mean vector of the current node data. Let x_{nf} be the mean vector of the fused data at the current node. Let x_{c1} be the mean vector of the target training class data. Let x_{c2}, x_{c3} be the mean vectors of the remaining training classes. A Euclidean distance ratio is defined as:

$$r_{dist} = d_{c1} / \min(d_{c2}, d_{c3}) , \qquad (7)$$

where d_{ci} is the Euclidean distance (4) between x_n and x_{ci}. We then use the following pseudocode to make our fusion decisions.

```
if (r_dist <= 1.0)
  fuse_4class = 1;   fuse_4carry = 1;
  class_ind = classify x_n;
  if (class_ind >= 70%)  check class_fuse;
  end
else
  fuse_4class = 0;   fuse_4carry = 0;
  if {(d_c1 <= 3s_c1 ) & (d_c2 <= 3s_c2 ) & (d_c2 <= 3s_c2 )}
    class_ind = classify x_n;
    if (class_ind == target class)  fuse_4class = 1;
      if (class_ind >= 70%)
        fuse_4carry = 1;
        class_fuse = classify x_nf;
        if (class_ind > class_fuse)
          class_fuse = class_ind;
        end
      end
    end
  else
    reject this data;
  end
end
```

There are two outcomes to the fusion decision. First we decide whether or not to fuse the data at the current node. If the current node has bad data, fusion can pull up the performance, however, we may not want to carry the bad data forward to the next node (the second fusion decision outcome). *fuse_4class* is a flag indicating whether or not to fuse for the current classification. *fuse_4carry* is a flag indicating whether or not to include data from the current node in the fused data that is carried forward. In Figures 4 and 5 we show the correct classification improvement gained by fusing from node to node for the acoustic and seismic sensors, respectively. For the acoustic sensor we show classification results from the AAV data, while using DW data for the seismic sensor results. In the case of the acoustic data, the mean correct classification performance across all nodes increases from 70% for independent operation to 93% with node to node fusion across the network. Similarly, the seismic correct classification performance increases from 42% to 52%.

4.1.3 Fusion between sensors

After fusion from node to node of the individual sensors, we look at the benefit of fusing the acoustic and seismic sensor data at the same node. To do so, we employ the following definitions. Let r_{dist} be defined as in (7) but with the new data types (*a* - acoustic, *s* - seismic, and *as* – a concatenated acoustic/seismic vector). Let x_a be the mean vector of the current node acoustic data after fusion from node to node. Let x_s be the mean vector of the current node seismic data after fusion from node to node. Let x_{as} = x_a concatenated with x_s (dumb fusion). Let x_{asf} = smart fusion of x_a with x_s. Let x_{in} be the data input to the classifier. Now, we employ two steps in the sensor fusion process as shown in the pseudocode below.
First we employ a smart sensor fusion routine:

```
indx = min(r_a_dist, r_s_dist, r_as_dist)
if (indx == 1)  x_in = x_a;
elseif (indx == 2)  x_in = x_s;
elseif (indx == 3)  x_in = x_as;
end
```

Next, we employ a final fusion routine:

```
class_acst = classify x_a;
class_seis = classify x_s;
class_as_dumb = classify x_as;
class_as_smart = classify x_asf;
if { (class_acst >= 70%) | (class_seis >= 70%) |
(class_as_ind >= 70%) }
  class_final_fuse = max (class_acst, class_seis,
class_as_dumb, class_as_smart)
end
```

Figure 6 shows the results of fusion at each stage. The classification performance is averaged over all the nodes for each vehicle class. The correct classification performance improves at each stage of fusion processing as shown in Table 1. The results indicate that the fusion based on value of information helps in improving the decision accuracy at each node significantly.

	AAV	DW	HMMV
acoustic independent	70 %	58 %	46%
seismic independent	72%	42 %	24%
Acoustic fusion	93%	80%	69%
seismic fusion	93%	52%	31%
acoustic and seismic, independent	76%	55%	58%
acoustic and seismic, with fusion	95%	90%	77 %

Table 1: Summary of classification performance

4.2 Detection experiments

For the detection experiments also both acoustic and seismic data were considered. First, only acoustic data from individual nodes were used. A threshold value was initially set which was varied adaptively based on the background energy. The power spectral density of acoustic data was computed using 1024 point FFT and it was downsampled by 8. The energy of the downsampled version of the power spectral density was computed. This energy was compared with the threshold value. If the energy was above the threshold value, it was decided that the target was detected. The time of detection and the confidence on detection were also calculated. The detection and time of detection were compared with the ground truth. If the target was detected when it is supposed to be and if the time of detection is within the region of interest then it was counted towards calculating the probability of detection. If the detection time is outside the region of interest (missed detection) and if a target was detected when it should not have been (false alarm) it was counted towards computing the probability of false alarm. The probability of detection and false alarm using only acoustic data from individual nodes without any fusion for AAV, DW and HMMWV are: 0.8824, 0.8677, 0.8382 & 0.1176, 0.1323, 0.1618, respectively. Similarly, the probability of detection and false alarm using only seismic data from individual nodes without any fusion for AAV, DW and HMMWV are: 0.8030, 0.7910, 0.5735 & 0.1970, 0.2090, 0.4265, respectively.

Next, the mutual information based value of information measure was used on the energy of power spectral density to make a decision of fusing data between sensors - acoustic and seismic on each individual node. The detector was tested using the fused data on each node. The probability of detection and false alarm were computed as described above. The probability of detection of this intelligently fused data for AAV, DW and HMMWV is: 0.9394, 0.9105 and 0.8529, respectively. The probability of false alarm is not provided here because it is equal to 1 – probability of detection since both false alarm and missed detections are combined together. These results are summarized in Figure 7 in the form of a bar graph. From this, it can be seen that the intelligent sensor data fusion based on value of information significantly improves the detection accuracy. This type of fusion especially helps in difficult data as in the case of HMMWV.

5 Conclusions

In this paper, we developed measures for value of information and we used these measures to make a decision of when to fuse information from neighboring nodes and between sensors. In the case of a detector we used this measure in the case of fusing data between sensors at present. In the future, we will also use these measures to make a decision of fusing information from neighboring nodes. In the case of classification, we have demonstrated that using these measures while fusing data between sensors and from neighboring nodes improve the classification accuracy significantly. However, our rejection algorithm is not robust at present. Our future work focuses on improving this algorithm. Our future work also focuses on further studying the mutual information based measure while fusing data between sensors and from node to node. In short, our results indicate that fusion based on value of information improves the decision accuracy – classification accuracy and probability of detection significantly.

6 References

1. S. Kadambe, ``Information theoretic based sensor discrimination for information fusion and cluster formation in a network of distributed sensors," in *Proc. of 4th annual conference on information fusion,*, Montreal, Quebec, Canada, August 7-10, 2001, pp. ThC1-19-ThC1-25.
2. S. Kadambe, ``Feature discovery and sensor discrimination in a network of distributed sensors for target tracking," in *Proc. of IEEE workshop on*

Statistical signal processing, Singapore, August 6-8, 2001, pp. 126-129.

3. R. Battti, “Using mutual information for selecting features in supervised neural net learning,” *IEEE Trans. On Neural Network*, vol. 5, no. 4, July 1994, pp. 537-550.
4. S. C. A. Thomopoulos, “Sensor selectivity and intelligent data fusion,” *Proc. Of the IEEE MIT’94*, October 2-5, 1994, Las Vegas, NV, pp. 529-537.
5. J. Manyika and H. Durrant-Whyte, Data fusion and sensor management: An information theoretic approach, Prentice Hall, 1994.
6. Papoulis, Probability, Random variables and Stochastic Processes, Second edition, McGraw Hill, 1984, pp. 500-567.
7. Y. Hen Wu, “Maximum likelihood based classifier,” *SensIT PI meeting*, Jan 15-17, 2002, Santa Fe, New Mexico.
8. S. Beck, “Energy based detector,” *SensIT PI meeting*, Jan 15-17, 2002, Santa Fe, New Mexico.
9. www.sensoria.com

Ackowledgement:
This work is funded by DARPA/IXO under the contract # F30602-01-C-0192. The authors would like to thank in particular Dr. Sri Kumar for funding this work under the SensIT program.

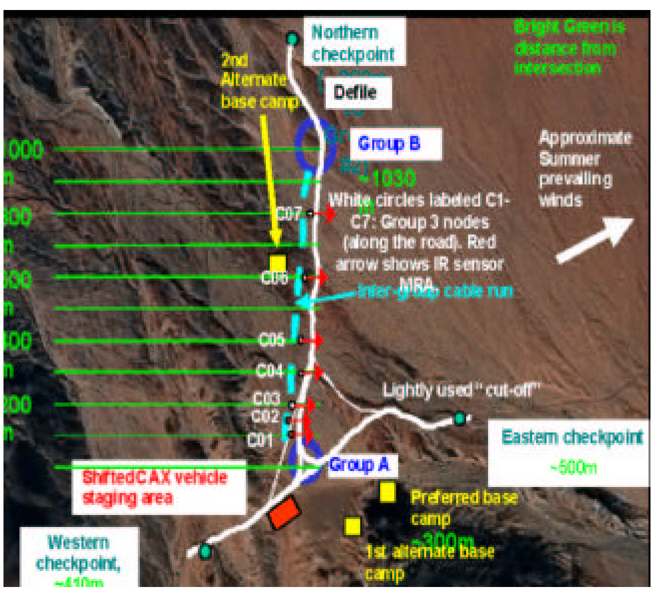

Figure 1: Sensor node distribution at Twenty nine Palms, CA

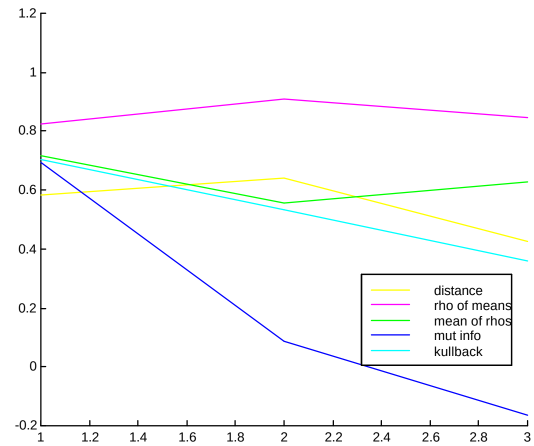

Figure 2: Performance correlation of previous node data

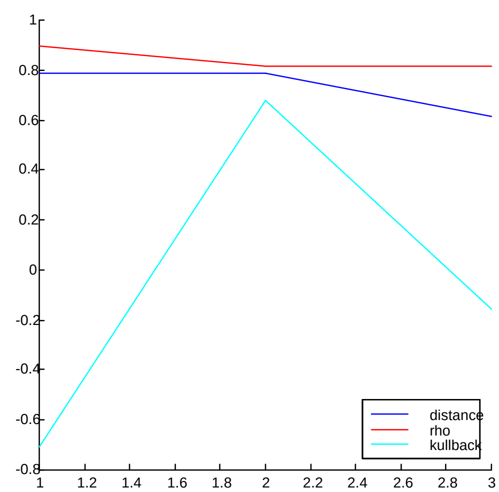

Figure 3: Performance correlation of training class data

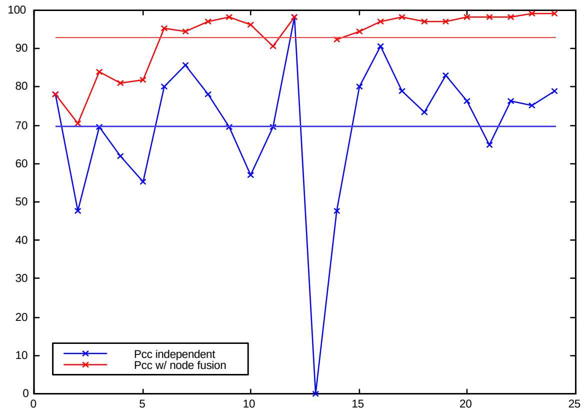

Figure 4: Performance of node fusion for the AAV with acoustic sensor data

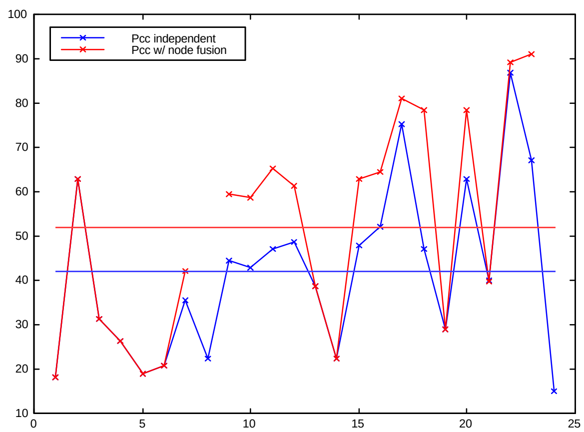

Figure 5: Performance of node fusion for the DW with seismic sensor data

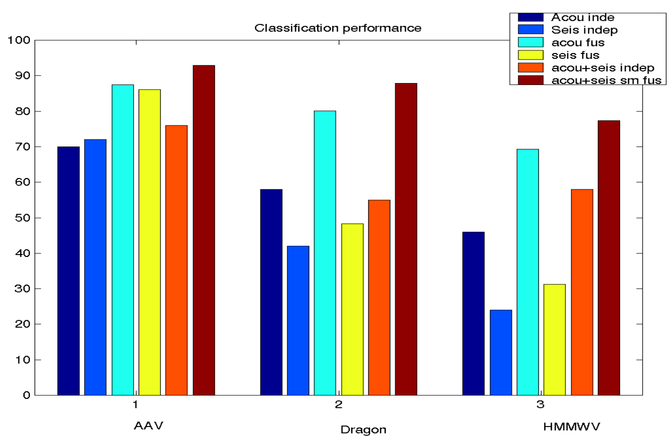

Figure 6: Average correct classification performance at each step in the fusion process

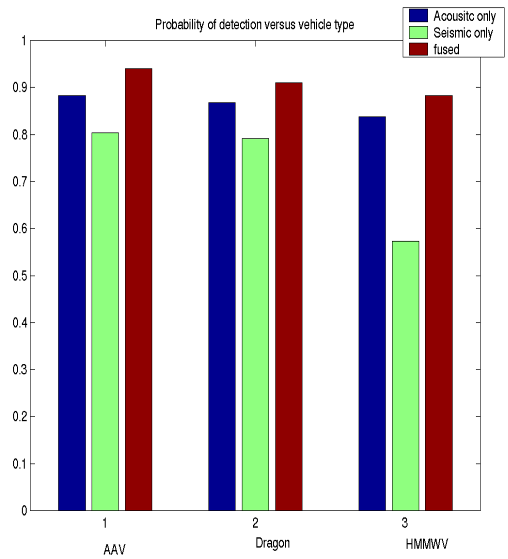

Figure 7: Performance of a detector